



UNITED STATES PATENT AND TRADEMARK OFFICE

15
UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/698,897	10/31/2003	Windsor Wee Sun Hsu	ARC920030068US1	2972
7590	12/29/2005		EXAMINER	
Mark C. McCabe IBM Corporation IP Law C4TA/J2B 650 Harry Road San Jose, CA 95120			MOORE, PATRICK M	
			ART UNIT	PAPER NUMBER
			2188	
			DATE MAILED: 12/29/2005	

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/698,897	HSU ET AL.	
	Examiner Patrick M. Moore	Art Unit 2188	

– The MAILING DATE of this communication appears on the cover sheet with the correspondence address –
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 31 October 2003.
 2a) This action is **FINAL**. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-18 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 1-18 is/are rejected.
 7) Claim(s) _____ is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
 Paper No(s)/Mail Date _____.

4) Interview Summary (PTO-413)
 Paper No(s)/Mail Date. _____.
 5) Notice of Informal Patent Application (PTO-152)
 6) Other: _____.

DETAILED ACTION

1. Claims 1-18 have been examined.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

2. Claims 1, 10-12 & 17 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.
 - a. Claim 1 recites the limitation "its replacement cost" in Line 3. It is unclear whether the term "its" refers to the replacement cost of the real cache partition or the data stored therein. In an effort to provide a complete examination on the merits, Examiner assumes Applicant is referring to the replacement cost of the data.
 - b. Claim 10 recites the limitation "the overall cost" in Lines 2 & 3 and the limitation "the cost" in Line 3. There is insufficient antecedent basis for these limitations in the claim.
 - c. Claim 11 recites the limitation "the hit/miss statistics" in Lines 1-2. There is insufficient antecedent basis for this limitation in the claim.
 - d. Claim 12 recites the limitation "the stack distance" in Line 1. There is insufficient antecedent basis for this limitation in the claim.

e. Claim 17 recites the limitation "the overall cost" twice in Line 2 and the limitation "the cost" in Line 3. There is insufficient antecedent basis for these limitations in the claim.

Appropriate correction is required.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 1-4, 6, 8-9, 13-15, 16 & 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Saulsbury et al. (US Patent # 5,900,011), herein Saulsbury, in view of Lamberts (US Patent # 6,418,510).

a. As per Claim 1, Saulsbury discloses a cost-adaptive cache, comprising: a partitioned real cache [Saulsbury: Figure 1, # 104 & Column 2, Lines 29-33]; and a partitioned phantom cache to provide a directory of information pertaining to blocks of data which do not qualify for inclusion in the real cache [Saulsbury: Figure 1, # 106 & Column 2, Lines 34-38], whereby the partitions in the phantom cache correspond to the partitions in the real cache [Saulsbury: Column 2, Lines 37-38]. Saulsbury defines a victim cache to be functionally identical to the "phantom cache" claimed by Applicant. However, Saulsbury does not expressly disclose using the replacement cost to prioritize data stored in a cache.

b. Lamberts teaches that data is stored in each the real cache partitions according to its replacement cost and the cost-adaptive cache maximizes performance in a system by preferentially caching data that is more costly to replace [*Lamberts*: Column 4, Lines 11-18]. Saulsbury and Lamberts are analogous art because they are from the same field of endeavor: optimizing cache memories to increase computer performance. At the time of invention, it would have been obvious to a person of ordinary skill in the art to modify the victim data cache of Saulsbury to include the cost function calculations taught by Lambert. The suggestion/motivation for doing so would have been to reduce the total access time for data stored in a computer system, which is explicitly stated by Lamberts in Column 4, Lines 29-37. Such motivation to combine applies to each of the subsequent dependent claims, below as Claims 2-15.

c. As per Claim 2, Lamberts further discloses the cost-adaptive cache of claim 1 wherein the real cache comprises a variable number of blocks for storing data [*Lamberts*: Column 4, Lines 61-68].

d. As per Claim 3, Lamberts further discloses the cost-adaptive cache of claim 1 wherein the real cache includes a configurable number of partitions [*Lamberts*: Column 6, Lines 55-663].

e. As per Claim 4, Lamberts further discloses the cost-adaptive cache of claim 3 wherein the configurable number of partitions are identified according to a replacement cost of data included within each of the partitions [*Lamberts*: Figure 3, # 60, 62 & Column 7, Lines 36-38].

- f. As per Claim 6, Saulsbury further discloses the cost-adaptive cache of claim 1 wherein the total size of corresponding partitions in the real cache and the phantom cache are less than or equal to the total size of the real cache [*Saulsbury*: Column 15, Lines 25-30].
- g. As per Claim 8, Lamberts discloses the cost-adaptive cache of claim 1 further comprises maintaining hit/miss statistics for each of the partitions in the real cache and each of the partitions in the phantom cache [*Lamberts*: Column 3, Lines 23-34].
- h. As per Claim 9, Lamberts discloses the cost-adaptive cache of claim 1 further comprises moving blocks between partitions within the real cache and the phantom cache in response to hits and misses in the caches [*Lamberts*: Column 3, Lines 35-37].
- i. As per Claim 13, Saulsbury further discloses the cost-adaptive cache of claim 9 wherein when a data block is evicted from the real cache, it is moved into the corresponding partition in the phantom cache [*Saulsbury*: Column 2, Lines 37-38 & Column 8, Lines 4-9].
- j. As per Claim 14, Saulsbury further discloses the cost-adaptive cache of claim 13 wherein a block of data can be evicted from the phantom cache in order to make room for the data block evicted from the real cache [*Saulsbury*: Column 3, Lines 6-12].
- k. As per Claim 15, Lamberts further discloses cost-adaptive cache of claim 1 wherein the replacement cost of a block of data is obtained by observing the

length of time needed to service a request for that data [*Lamberts*: Column 7, Lines 19-29].

I. As per Claim 16, Saulsbury discloses a method for dynamically partitioning a storage system cache according to a replacement cost associated with data stored in the cache, the cache holding the data as blocks of data, the method comprising the steps of: maintaining a history of recently evicted data blocks for each partition [*Saulsbury*: Column 8, Lines 4-9]; and *Lamberts* discloses assigning data to one partition based on a cost associated with not keeping the data in the cache [*Lamberts*: Column 7, Lines 19-29]; determining a future size for each partition based on the history and the cost associated with not keeping the data in the cache [*Lamberts*: Column 7, Lines 19-29]; and whereby the cache's performance is dynamically maximized by preferentially caching data that are most costly to replace [*Lamberts*: Column 4, Lines 29-37]. As for Claims 16-18, identical motivation to combine Saulsbury with *Lamberts* exists as noted above for Claims 1-15.

m. As per Claim 18, *Lamberts* further discloses the method of claim 16 wherein the cost of not keeping the data in the cache is obtained by observing the length of time needed to service a request for that data [*Lamberts*: Column 7, Lines 19-29].

4. Claims 5, 7, 10-12 & 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Saulsbury (US Patent # 5,900,011) and *Lamberts* (US Patent #

6,418,510) as applied to Claims 1-4, 6, 8-9, 13-15, 16 & 18 above, and further in view of Smith (US Patent # 5,394,531).

- a. Saulsbury and Lamberts do not expressly disclose the using the replacement cost to define partition size in a cache, but Smith teaches the cost-adaptive cache of claim 3 wherein the partitions each have a pair of associated replacement cost values which define the boundaries for each of the partitions [*Smith: Column 2, Lines 62-68*]. Note that Smith refers to the partition's cost value as a hit/miss ratio, which is based on the number of times a data block is requested.
- b. Furthermore, Saulsbury, Lamberts and Smith are analogous art because they are all from similar problem solving areas: increasing computing efficiency through the optimized use of cache memories. At the time of invention, it would have been obvious to a person of ordinary skill in the art to modify the victim data cache of Saulsbury, combined with the cost adaptive techniques of Lamberts, by using the dynamic partition segmentation disclosed by Smith. The suggestion/motivation for doing so would have been to increase the hit/miss ratio of the data stored in a cache memory by adjusting the cache partition sizes as is directly disclosed by Smith in Column 2, Lines 54-61. Finally, identical motivation to combine applies to each of the subsequent claims, below as Claims 7, 10-12 & 17.
- c. As per Claim 7, Smith further discloses the cost-adaptive cache of claim 1 wherein a target size is associated with each of the partitions in the real cache

and the target sizes can be fixed, dynamic or adjusted periodically [*Smith*: Column 2, Lines 62-64].

d. As per Claim 10, Smith further discloses the cost-adaptive cache of claim 9 further comprises adjusting the sizes of the partitions in the real cache to minimize the overall cost of servicing data requests [*Smith*: Column 3, Lines 43-53], and Lamberts discloses wherein the overall cost comprises the number of times data is requested and the cost of satisfying each of the requests [*Lamberts*: Column 7, Lines 19-29].

e. As per Claim 11, Lamberts further discloses the cost-adaptive cache of claim 10 wherein the adjustment [*Lamberts*: Column 7, Lines 51-52] is based on the hit/miss statistics of the partitions [*Lamberts*: Column 3, Lines 35-37] and the relative replacement cost of blocks in the partitions [*Lamberts*: Column 7, Lines 19-29].

f. As per Claim 12, Lamberts further discloses the cost-adaptive cache of claim 10 wherein the adjustment is based on the stack distance of a hit in the phantom cache [*Lamberts*: Column 3, Lines 42-49].

g. As per Claim 17, Smith further discloses the method of claim 16 wherein the future size of each partition is determined so as to minimize the overall cost of servicing requests for data [*Smith*: Column 3, Lines 43-53], and Lamberts discloses that the overall cost comprises the number of times the data is requested and the cost of satisfying each of the requests [*Lamberts*: Column 7, Lines 19-29].

Conclusion

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Doren et al. (US Patent # 6,101,581) discloses maintaining a victim cache until the CPU no longer needs the associated data. Lewis et al. (US PGPub # 2002/0073283) discloses target partition sizes and configurable and dynamic partitions. Olarig et al. (US PGPub # 2003/0065886) discloses dynamic cache partitioning with adjustable sizing.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Patrick M. Moore whose telephone number is (571) 272-1239. The examiner can normally be reached on M-F 8:30AM - 5:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Mano Padmanabahn can be reached on (571) 272-4210. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



MANO PADMANABHAN
SUPERVISORY PATENT EXAMINER

Application/Control Number: 10/698,897
Art Unit: 2188

Page 10

PMM